Kwangho Kim

Room 409, Political Science and Economics Hall, Korea University
145 Anam-ro, Seongbuk-gu, Seoul 02841, Republic of Korea
kwanghk@korea.ac.kr ▪ +82 10-4107-9265 · +1 (650) 285-7258
Google Scholar ▪ ORCID 0000-0001-8846-363X

Last updated: August 2026

RESEARCH INTERESTS

Causal and counterfactual learning; semiparametric and nonparametric statistics; applied algebraic topology; fair and robust machine learning; applications to healthcare policy, finance, manufacturing, and biomedical research.

ACADEMIC APPOINTMENTS

2023–Present **Korea University** *Seoul, Korea*
Associate Professor (2026–Present); Assistant Professor (2023–2026)
- Department of Statistics, College of Political Science and Economics
- Department of Data Science, College of Informatics

EDUCATION

2020 **Carnegie Mellon University** *Pittsburgh, PA, United States*
Joint Ph.D. in Statistics and Machine Learning
- Advisor: Edward H. Kennedy Co-advisor: Larry A. Wasserman

2018 **Carnegie Mellon University** *Pittsburgh, PA, United States*
M.S. in Machine Learning

2016 **Stanford University** *Stanford, CA, United States*
M.S. in Statistics

2012 **Korea Advanced Institute of Science and Technology (KAIST)** *Daejeon, Korea*
B.S. in Electrical Engineering; B.S. in Mathematical Sciences
- Graduated *summa cum laude*

PROFESSIONAL EXPERIENCE

2020–2023 **Harvard Medical School** *Boston, MA, United States*
Marshall J. Seidman Postdoctoral Fellow, Department of Health Care Policy
- Research supervisor: José R. Zubizarreta

2019 **Google** *New York, NY, United States*
Data Scientist Intern, Advanced Measurement Technology

2015–2016 **Stanford Graduate School of Business** *Stanford, CA, United States*
Graduate Research Assistant
- Research supervisor: Guido W. Imbens

2013–2014 **Samsung Asset Management** *Seoul, Korea*
Quantitative Researcher

2011 **Citibank** *Seoul, Korea*
Quantitative Analyst Intern

PUBLICATIONS

[*] *Corresponding author.* [†] *Student advisee.*

Refereed Journal Articles

[J8] JungHyun Kim[†], Dongwook Lee, and **Kwangho Kim**[*] (2026). Causal evaluation of early-stage fast-charging protocols under cell-to-cell variability. *IEEE Transactions on Industrial Informatics*.

[J7] Youmi Suk, Chan Park, Chenguang Pan, and **Kwangho Kim**[*] (2026). Fair and robust estimation of heterogeneous treatment effects for optimal policies in multilevel studies. *Multivariate Behavioral Research*, pp.1-20.

[J6] **Kwangho Kim**[*], Jisu Kim, and Edward H. Kennedy (2026). Causal k-means clustering. *Journal of the Royal Statistical Society, Series B*. (2020 WNAR Student Paper Award)

[J5] Munil Yang, **Kwangho Kim**, and Dongha Kim (2025). From prediction to action: A framework for business closure analysis and LLM-based consulting strategy generation. *Journal of the Korean Statistical Society*, pp. 1–19.

[J4] Alloyssius E. G. Gorospe, Dongwoo Kang, Kyunghyun Kim, **Kwangho Kim**, and Dongwook Lee (2025). Ultralong cycle lifespan in lithium metal batteries unlocked by a lithiophilic montmorillonite separator. *Materials Letters*, Article 139240.

[J3] **Kwangho Kim**, Bijan A. Niknam, and José R. Zubizarreta (2024). Scalable kernel balancing weights in a nationwide observational study of hospital profit status and heart attack outcomes. *Biostatistics*, 25(3), pp. 736–753.

[J2] Jacqueline E. Rudolph, **Kwangho Kim**, Edward H. Kennedy, and Ashley I. Naimi (2023). Estimation of the time-varying incremental effect of low-dose aspirin on incidence of pregnancy. *Epidemiology*, 34(1), pp. 38–44.

[J1] **Kwangho Kim**[*], Edward H. Kennedy, and Ashley I. Naimi (2021). Incremental intervention effects in studies with many timepoints and dropout. *Journal of Causal Inference*, 9(1), pp. 302–344. (2020 ASA Student Paper Award, Nonparametric Statistics Section)

Refereed Conference Proceedings

[C7] Hajin Lee[†] and **Kwangho Kim**[*] (2026). Provably correct k-means clustering of persistence diagrams. In *Proceedings of the 42nd Conference on Uncertainty in Artificial Intelligence (UAI)*.

[C6] **Kwangho Kim** and Jisu Kim (2026). Semi-supervised learning with noisy proxy covariates: Generalization bounds and distribution regression. In *Proceedings of the 43rd International Conference on Machine Learning (ICML)*.

[C5] **Kwangho Kim**[*] and Hajin Lee[†] (2026). Topological causal effects. In *Proceedings of the International Conference on Learning Representations (ICLR)*.

[C4] **Kwangho Kim**[*], Jisu Kim, Larry A. Wasserman, and Edward H. Kennedy (2024). Hierarchical and density-based causal clustering. In *Advances in Neural Information Processing Systems (NeurIPS) 37*, pp. 30363–30393.

[C3] **Kwangho Kim** and José R. Zubizarreta (2023). Fair and robust estimation of heterogeneous treatment effects for policy learning. In *Proceedings of the 40th International Conference on Machine Learn-*

ing (ICML), pp. 16997–17014.

[C2] **Kwangho Kim**, Edward H. Kennedy, and José R. Zubizarreta (2022). Doubly robust counterfactual classification. In *Advances in Neural Information Processing Systems (NeurIPS) 35*, pp. 34831–34845.

[C1] **Kwangho Kim**[*], Jisu Kim, Manzil Zaheer, Joon-sik Kim, Larry A. Wasserman, and Frédéric Chazal (2020). PLLay: Efficient topological layer based on persistent landscapes. In *Advances in Neural Information Processing Systems (NeurIPS) 33*, pp. 15965–15977.

MANUSCRIPTS UNDER REVIEW OR IN REVISION

[†] *Student advisee.*

[M17] Hajin Lee[†] and **Kwangho Kim**. Will our treatment policy really work? Multiscale segmentation of treatment benefit regions.

[M16] Hajin Lee[†], Woohyeok Yoo[†], and **Kwangho Kim**. Q-Learning with Buckley–James Imputation under Right Censoring.

[M15] Uhjin Shin[†] and **Kwangho Kim**. Learning under counterfactual shifts in image acquisition.

[M14] JungHyun Kim[†] and **Kwangho Kim**. Causal Clustering for Subgroup Discovery under Longitudinal Treatment Regimes.

[M13] Uhjin Shin[†], Eun Kang, Kyounghae Kim, and **Kwangho Kim**. Sustained Engagement Strategies for Chronic Low Back Pain Using PROPEL Trial Data.

[M12] **Kwangho Kim**. Orthogonal masked counterfactual learning from incomplete data.

[M11] Hajin Lee[†], Jisu Kim, and **Kwangho Kim**. ECLayr: A Fast and Robust Topological Layer via Euler Characteristic Curves.

[M10] **Kwangho Kim** and Jos'e R. Zubizarreta. Bias-controlled learning of treatment-responsive subgroups.

[M9] **Kwangho Kim**. Counterfactual risk minimization: Learning and optimization under interventions.

[M8] Woohyeok Yoo[†], Hajin Lee[†], and **Kwangho Kim**. Doubly Robust Q-Learning for Right-censored Survival Outcomes via AIPCW Temporal-Difference Targets.

[M7] Woohyeok Yoo[†], Uhjin Shin[†], Mingue Park, and **Kwangho Kim**. Profile-based Kernel Balancing for Generalization and Transportability.

[M6] **Kwangho Kim**, Hajin Lee[†], Haegang Lee, Heebum Lee, Jisu Kim, and Alessandro Rinaldo. Topology-regularized signal synthesis under distribution shift.

[M5] **Kwangho Kim**. Data-poisoning audits for causal effect estimation. arXiv:2607.19692.

[M4] **Kwangho Kim**. Geometry-adaptive counterfactual distribution learning with diffusion-guided smoothing. arXiv:2605.25811.

[M3] **Kwangho Kim**. Counterfactual regression under intervention-driven shift. arXiv:2504.02694.

[M2] **Kwangho Kim**, Alan Mishler, and Jos'e R. Zubizarreta. Counterfactual mean-variance optimization. arXiv:2209.09538. *In revision, Journal of Business & Economic Statistics*.

[M1] **Kwangho Kim**, Jisu Kim, and Edward H. Kennedy. Causal effects based on distributional distances. arXiv:1806.02935.

PATENTS

Granted

2025 **Method for reconstructing financial instruments or indices into a virtual index composed of investable products, and a customized index-based investment product platform driven by investor demand** *KR 10-2767926*
금융투자상품 또는 금융 지수를 투자 가능 금융 상품으로 조합된 가상의 지수로 재구성하여 운용하는 방법 및 수요자 투자 요구에 기반한 맞춤형 지수 기반 투자 상품 운용 플랫폼

2025 **Method for operating a virtual fund that replicates a non-investable investment fund through combinations of investable products** *KR 10-2767920*
투자 가능한 상품의 조합으로 투자 불가능한 투자 펀드를 모방한 가상 펀드 운영 방법

2024 **Method for stepwise replication of time-series data characteristics through combinations of basis time series, and computer apparatus therefor** *KR 10-2591935*
기저 시계열의 조합을 통한 시계열 데이터의 단계적 특성 모사 방법 및 이를 수행하는 컴퓨터 장치

Pending Applications

2025 **Apparatus for performing generalization and transportation estimation using profile-based kernel balancing, and method thereof** *KR 10-2025-0212682*
프로파일 기반 커널 밸런싱을 이용한 일반화 및 전이 추정을 수행하는 장치 및 그 동작 방법

2024 **Apparatus and method for constructing a minimum-cost financial portfolio replicating the characteristics of a target asset under resource constraints** *KR 10-2024-0166773*
제한된 자원 하에서 최소 비용으로 타겟 자산의 특성을 모사하는 금융 포트폴리오를 구축하는 장치, 방법 및 기록매체

RESEARCH GRANTS

2026–2029 **PI**, Samsung Science and Technology Foundation *SSTF-BA2502-01*

2025–2026 **PI**, Early-Career Researcher Program (Seed) *RS-2025-24534596*
National Research Foundation of Korea (NRF)

2024–2027 **Co-I**, Mid-Career Researcher Program *RS-2024-00335008*
National Research Foundation of Korea (NRF)

HONORS AND AWARDS

2024–2025 **Excellence in Teaching Award** *Korea University*
- 2025 — Introduction to Causal Inference (STAT 407)
- 2024 — Introduction to Mathematical Statistics (STAT 232)

2020 **Best Student Paper Award** *International Biometric Society, WNAR*

2020 **IMS Hannan Graduate Student Travel Award** *Institute of Mathematical Statistics*

2020 **ASA Student Paper Award** *American Statistical Association*
Nonparametric Statistics Section

2017 **CMU Citadel Datathon, Top Prize** *Citadel & Citadel Securities*

2014	**Korean Government Scholarship**	*Ministry of Education, Korea*
2011	**Finance Research Paper Competition, Top Prize**	*Citibank & Korea Institute of Finance*
2010	**Statistics Application Research Award**	*Statistics Korea*
2009	**Honor Student** *Academic Excellence Award*	*KAIST*
2008	**KAIST President's Award** *For organizing and leading the undergraduate dormitory residents' council*	*KAIST*
2006	**Presidential Science Scholarship**	*Korea Science and Engineering Foundation*

SELECTED TALKS AND PRESENTATIONS

Invited Talks

2026	**7th IMS Asia Pacific Rim Meeting (IMS-APRM)** *"Efficient counterfactual distribution learning via geometric adaptation"*	*Hong Kong*
2025	**Taipei International Statistical Symposium and 13th ICSA International Conference** *"Toward flexible and efficient counterfactual density estimation"*	*Taiwan*
2024	**International Conference for Statistics and Data Science (ICSDS)** *"Counterfactual prediction"*	*Taiwan*
2024	**The Institute of Statistical Mathematics** *"Causal and counterfactual learning"*	*Japan*
2024	**7th International Conference on Econometrics and Statistics (EcoSta)** *"Counterfactual prediction"*	*China*
2023	**6th International Conference on Econometrics and Statistics (EcoSta)** *"Causal clustering"*	*Japan*
2022	**Korean International Statistical Society (KISS)** *"Scalable kernel balancing weights: Application to healthcare policy"*	*United States*
2021	**The Korean Statistical Society** *"Causal k-means clustering"*	*Korea*
2019	**Google Research** *"Incremental intervention effects in studies with many timepoints and dropout"*	*United States*

Contributed Talks and Posters

2026	**43rd International Conference on Machine Learning (ICML)** *Poster — "Semi-supervised learning with noisy proxy covariates: Generalization bounds and distribution regression"*	*Korea*
2024	**38th Conference on Neural Information Processing Systems (NeurIPS)** *Poster — "Hierarchical and density-based causal clustering"*	*Canada*
2023	**40th International Conference on Machine Learning (ICML)** *Poster — "Fair and robust estimation of heterogeneous treatment effects for policy learning"*	*United States*

2022 **36th Conference on Neural Information Processing Systems (NeurIPS)** *United States*
Poster — "Doubly robust counterfactual classification"

2020 **34th Conference on Neural Information Processing Systems (NeurIPS)** *United States*
Poster — "PLLay: Efficient topological layer based on persistent landscapes"

2020 **Joint Statistical Meetings (JSM)** *United States*
Contributed paper — "Incremental intervention effects in studies with many timepoints and dropout"

2020 **International Conference on Health Policy Statistics (ICHPS)** *United States*
Contributed paper — "Causal clustering: A new framework for analysis of treatment effect heterogeneity"

2018 **Atlantic Causal Inference Conference (ACIC)** *United States*
Poster — "Causal effects based on distributional distances"

TEACHING AND ADVISING

2023–Present **Korea University** *Seoul, Korea*
Instructor, Department of Statistics and Department of Data Science

- **STAT 221** — *Introduction to Probability Theory*
- **STAT 232** — *Introduction to Mathematical Statistics*
- **STAT 407** — *Introduction to Causal Inference*
- **STAT 640** — *Modern Causal Inference*
- **STAT 720** — *Fundamentals of Data Science*
- **IDS 08** — *Introduction to Interpretable Machine Learning*

2016–2020 **Carnegie Mellon University** *Pittsburgh, PA, United States*
Teaching Assistant, Department of Statistics and Data Science

- **36-219** — *Probability Theory and Random Processes*
- **36-350** — *Statistical Computing*
- **36-402** — *Advanced Data Analysis*
- **46-921** — *Financial Data Science*
- **46-926** — *Statistical and Machine Learning*
- **46-929** — *Financial Time Series Analysis*

2018 **Carnegie Mellon University** *Pittsburgh, PA, United States*
Research Supervisor, Summer Undergraduate Research Program

2015 **Stanford University** *Stanford, CA, United States*
Statistical Consultant, Department of Statistics

SOFTWARE

- **sbw** — R package implementing the stable balancing weights method of Zubizarreta (2015). Available on CRAN.

REFEREE SERVICE

Journals (selected)

- Annals of Applied Statistics
- Journal of the Royal Statistical Society, Series B
- Annals of Statistics
- Biostatistics
- International Journal of Biostatistics
- Computational Statistics & Data Analysis
- Journal of Causal Inference
- Communications AI & Computing

Conferences (selected)

- Advances in Neural Information Processing Systems (NeurIPS)
- International Conference on Learning Representations (ICLR)
- International Conference on Machine Learning (ICML)

COMMITTEES AND OTHER SERVICE

- **Graduate Admissions Committee** (2023–2026). Department of Statistics, Korea University
- **Graduate Admissions Committee** (2024–2026). Graduate School of Interdisciplinary Data Science, Korea University
- **Ph.D. Admissions Committee** (2020). Machine Learning Department, Carnegie Mellon University
- **Seminar Coordinator** (2020). Topics in Causal Inference, Harvard University
- **Seminar Coordinator** (2019). Reading Group in Statistical Machine Learning, Carnegie Mellon University
- **Seminar Coordinator** (2019). Reading Group in Topological Data Analysis, Carnegie Mellon University

PROFESSIONAL CERTIFICATIONS

- **Certified Investment Manager** *Korea Financial Investment Association*
- **Certified Derivatives Investment Advisor** *Korea Financial Investment Association*